

OIPE

#2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/938,885

DATE: 09/26/2001

TIME: 12:07:56

Input Set : N:\Crf3\RULE60\09938885.txt

Output Set: N:\CRF3\09262001\I938885.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Bandman, Olga
8 Goli, Surya K.
9 Murry, Lynn E.

C--> 11 (ii) TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT

13 (iii) NUMBER OF SEQUENCES: 5

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
17 (B) STREET: 3174 Porter Drive
18 (C) CITY: Palo Alto
19 (D) STATE: CA
20 (E) COUNTRY: US
21 (F) ZIP: 94304

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Diskette
25 (B) COMPUTER: IBM Compatible
26 (C) OPERATING SYSTEM: DOS
27 (D) SOFTWARE: FastSEQ Version 2.0

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/938,885

C--> 31 (B) FILING DATE: 24-Aug-2001

32 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: 08/760,745
36 (B) FILING DATE: 05-DEC-1996

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Billings, Lucy J.
40 (B) REGISTRATION NUMBER: 36,749
41 (C) REFERENCE/DOCKET NUMBER: PF-0169 US

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: 415-855-0555
45 (B) TELEFAX: 415-845-4166
46 (C) TELEX:

49 (2) INFORMATION FOR SEQ ID NO: 1:

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 235 amino acids
53 (B) TYPE: amino acid
54 (C) STRANDEDNESS: single
55 (D) TOPOLOGY: linear

57 (vii) IMMEDIATE SOURCE:

58 (A) LIBRARY: lungast01
59 (B) CLONE: 876242

61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

63 Met Pro His Ala Phe Lys Pro Gly Asp Leu Val Phe Ala Lys Met Lys
64 1 5 10 15
65 Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Asp Ile Ala Asp Gly Ala

ENTERED

RAW SEQUENCE LISTING

DATE: 09/26/2001

PATENT APPLICATION: US/09/938,885

TIME: 12:07:56

Input Set : N:\Crf3\RULE60\09938885.txt

Output Set: N:\CRF3\09262001\I938885.raw

```

66          20          25          30
67 Val Lys Pro Pro Pro Asn Lys Tyr Pro Ile Phe Phe Phe Gly Thr His
68          35          40          45
69 Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe Pro Tyr Glu Glu Ser
70          50          55          60
71 Lys Glu Lys Phe Gly Lys Pro Asn Lys Arg Lys Gly Phe Ser Glu Gly
72          65          70          75          80
73 Leu Trp Glu Ile Glu Asn Asn Pro Thr Val Lys Ala Ser Gly Tyr Gln
74          85          90          95
75 Ser Ser Gln Lys Lys Ser Cys Val Glu Glu Pro Glu Pro Glu Pro Glu
76          100          105          110
77 Ala Ala Glu Gly Asp Gly Asp Lys Lys Gly Asn Ala Glu Gly Ser Ser
78          115          120          125
79 Asp Glu Glu Gly Lys Leu Val Ile Asp Glu Pro Ala Lys Glu Lys Asn
80          130          135          140
81 Glu Lys Gly Ala Leu Lys Arg Arg Ala Gly Asp Leu Leu Glu Asp Ser
82          145          150          155          160
83 Pro Lys Arg Pro Lys Glu Ala Glu Asn Pro Glu Gly Glu Glu Lys Glu
84          165          170          175
85 Ala Ala Thr Leu Glu Val Glu Arg Pro Leu Pro Met Glu Val Glu Lys
86          180          185          190
W--> 87 Asn Ser Thr Pro Ser Glu Pro Gly Ser Gly Arg Gly Pro Pro Xaa Xaa
88          195          200          205
W--> 89 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Glu Glu Ala Thr Lys Glu Asp Ala
90          210          215          220
91 Glu Ala Pro Gly Ile Lys Ser His Glu Ser Leu
92          225          230          235
94 (2) INFORMATION FOR SEQ ID NO: 2:
96     (i) SEQUENCE CHARACTERISTICS:
97         (A) LENGTH: 869 base pairs
98         (B) TYPE: nucleic acid
99         (C) STRANDEDNESS: single
100        (D) TOPOLOGY: linear
102     (vii) IMMEDIATE SOURCE:
103         (A) LIBRARY: lungast01
104         (B) CLONE: 876242
106     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
108 CCGCTGCAGC CGCTTTCTGC GGCCTGGGCC TCTCGCCGTC AGCATGCCAC ACGCCTTCAA      60
109 GCCCGGGGAC TTGGTGTTCG CTAAGATGAA GGGCTACCCT CACTGGCCTG CCAGGATCGA      120
110 CGACATCGCG GATGGCGCCG TGAAGCCCCC ACCCAACAAG TACCCCATCT TTTTCTTTGG      180
111 CACACACGAA ACAGCCTTCC TGGGCCCCAA AGACCTCTTC CTTACGAGG AATCCAAGGA      240
112 GAAGTTTGGC AAGCCCAACA AGAGGAAAGG GTTCAGCGAG GGGCTGTGGG AGATCGAGAA      300
113 CAACCCTACT GTCAAGGCTT CCGGCTATCA GTCCTCCCAG AAAAAGAGCT GTGTGGAAGA      360
114 GCCTGAACCA GAGCCCGAAG CTGCAGAGGG TGACGGTGAT AAGAAGGGGA ATGCAGAGGG      420
115 CAGCAGCGAC GAGGAAGGGA AGCTGGTCAT TGATGAGCCA GCCAAGGAGA AGAACGAGAA      480
116 AGGAGCGTTG AAGAGGAGAG CAGGGGACTT GCTGGAGGAC TCTCCTAAAC GTCCCAAGGA      540
117 GGCAGAAAC CCTGAAGGAG AGGAGAAGGA GGCAGCCACC TTGGAGGTTG AGAGGCCCTT      600
118 TCCTATGGAG GTGGAAGAAGA ATAGCACCCC CTCTGAGCCC GGCTCTGGCC GGGGGCCTCC      660
119 CCNNNNNNNN NNNNNNNNNN NNNNNNNNNA GGAAGAGGCT ACCAAGGAAG ATGCTGAGGC      720

```

RAW SEQUENCE LISTING

DATE: 09/26/2001

PATENT APPLICATION: US/09/938,885

TIME: 12:07:56

Input Set : N:\Crif3\RULE60\09938885.txt

Output Set: N:\CRF3\09262001\I938885.raw

```

120 CCCAGGCATC AAGAGTCATG AGAGCCTGTA GCCACCAATG TTTCAAGAGG AGCCCCCACC 780
121 CTGTTTCCTGC TGCTGTCTGG GTGCTACTGG GGAAACTGGC CATGGGCTGC AAAGTGGGNA 840
122 CCCCTTTTCC ANCNCANCT GNTNTTCTT 869

```

124 (2) INFORMATION FOR SEQ ID NO: 3:

126 (i) SEQUENCE CHARACTERISTICS:

127 (A) LENGTH: 240 amino acids

128 (B) TYPE: amino acid

129 (C) STRANDEDNESS: single

130 (D) TOPOLOGY: linear

132 (vii) IMMEDIATE SOURCE:

133 (A) LIBRARY: GenBank

134 (B) CLONE: 598956

136 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

138 Met Ser Arg Ser Asn Arg Gln Lys Glu Tyr Lys Cys Gly Asp Leu Val
139 1 5 10 15
140 Phe Ala Lys Met Lys Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Glu
141 20 25 30
142 Met Pro Glu Ala Ala Val Lys Ser Thr Ala Asn Lys Tyr Gln Val Phe
143 35 40 45
144 Phe Phe Gly Thr His Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe
145 50 55 60
146 Pro Tyr Glu Glu Ser Lys Glu Lys Phe Gly Lys Pro Asn Lys Arg Lys
147 65 70 75 80
148 Gly Phe Ser Glu Gly Leu Trp Glu Ile Glu Asn Asn Pro Thr Val Lys
149 85 90 95
150 Ala Ser Gly Tyr Gln Ser Ser Gln Lys Lys Ser Cys Val Glu Glu Pro
151 100 105 110
152 Glu Pro Glu Pro Glu Ala Ala Glu Gly Asp Gly Asp Lys Lys Gly Asn
153 115 120 125
154 Ala Glu Gly Ser Ser Asp Glu Glu Gly Lys Leu Val Ile Asp Glu Pro
155 130 135 140
156 Ala Lys Glu Lys Asn Glu Lys Gly Ala Leu Lys Arg Arg Ala Gly Asp
157 145 150 155 160
158 Leu Leu Glu Asp Ser Pro Lys Arg Pro Lys Glu Ala Glu Asn Pro Glu
159 165 170 175
160 Gly Glu Glu Lys Glu Ala Ala Thr Leu Glu Val Glu Arg Pro Leu Pro
161 180 185 190
162 Met Glu Val Glu Lys Asn Ser Thr Pro Ser Glu Pro Gly Ser Gly Arg
163 195 200 205
164 Gly Pro Pro Gln Glu Glu Glu Glu Glu Asp Glu Glu Glu Glu Ala
165 210 215 220
166 Thr Lys Glu Asp Ala Glu Ala Pro Gly Ile Arg Asp His Glu Ser Leu
167 225 230 235 240

```

169 (2) INFORMATION FOR SEQ ID NO: 4:

171 (i) SEQUENCE CHARACTERISTICS:

172 (A) LENGTH: 2376 base pairs

173 (B) TYPE: nucleic acid

174 (C) STRANDEDNESS: single

175 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING

DATE: 09/26/2001

PATENT APPLICATION: US/09/938,885

TIME: 12:07:56

Input Set : N:\Crif3\RULE60\09938885.txt

Output Set: N:\CRF3\09262001\I938885.raw

```

177 (vii) IMMEDIATE SOURCE:
178 (A) LIBRARY: GenBank
179 (B) CLONE: 598956
181 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
183 GAGGAGGAGT GGGGACCGGG CGGGGGGTGG AGGAAGAGGC CTCGCGCAGA GGAGGGAGCA 60
184 ATTGAATTC AAACACAAAC AACTCGACGA GCGCGCACCC ACCGCGCCGG AGCCTTGCCC 120
185 CGATCCGCGC CCGCCCCGTC CGTGCGGCGC GCGGGCGGAG ACGCCGTGGC CGCGCCGGAG 180
186 CTCGGGCCGG GGGCCACCAT CGAGGCGGGG GCCGCGCGAG GGCCGGAGCG GAGCGGCGCC 240
187 GCCACCGCCG CACGCGCAA CTTGGGCTCG CGCTTCCCGG CCCGGCGCGG AGCCCGGGGC 300
188 GCCCGGAGCC CCGCCATGTC GCGATCCAAC CGGCAGAAGG AGTACAAATG CGGGGACCTG 360
189 GTGTTCGCCA AGATGAAGGG CTACCCACAC TGGCCGGCCC GGATTGACGA GATGCCTGAG 420
190 GCTGCCGTGA AATCAACAGC CAACAAATAC CAAGTCTTTT TTTTCGGGAC CCACGAGACG 480
191 GCATTCCTGG GCCCAAAGA CCTCTTCCCT TACGAGGAAT CCAAGGAGAA GTTTGGCAAG 540
192 CCCAACAAGA GGAAAGGGT CAGCGAGGGG CTGTGGGAGA TCGAGAACAA CCCTACTGTC 600
193 AAGGCTCCG GCTATCAGTC CTCCAGAAA AAGAGCTGTG TGGAAGAGCC TGAACCAGAG 660
194 CCCGAAGCTG CAGAGGGTGA CGGTGATAAG AAGGGGAATG CAGAGGGCAG CAGCGACGAG 720
195 GAAGGGAAGC TGGTCATTGA TGAGCCAGCC AAGGAGAAGA ACGAGAAAGG AGCGTTGAAG 780
196 AGGAGAGCAG GGGACTTGCT GGAGACTCT CCTAAACGTC CCAAGGAGGC AGAAAACCT 840
197 GAAGGAGAGG AGAAGGAGGC AGCCACCTTG GAGGTTGAGA GGCCCTTCC TATGGAGGTG 900
198 GAAAAGAATA GCACCCCTC TGAGCCCGGC TCTGGCCGGG GGCCTCCCA AGAGGAAGAA 960
199 GAAGAGGAGG ATGAAGAGGA AGAGGCTACC AAGGAAGATG CTGAGGCCCC AGGCATCAGA 1020
200 GATCATGAGA GCCTGTAGCC ACCAATGTTT CAAGAGGAGC CCCCACCCTG TTCCTGCTGC 1080
201 TGTCTGGGTG CTACTGGGGA AACTGGCCAT GGCCTGCAA CTGGGAACCC CTTTCCACC 1140
202 CCAACCTGCT CTCCTCTTCT ACTCACTTTT CCCACTCAA GCCCAGCCCA TGGAGATTGA 1200
203 CCTGGATGGG GCAGGCCACC TGGCTCTCAC CTCTAGGTCC CCATACTCCT ATGATCTGAG 1260
204 TCAGAGCCAT GTCTTCTCCC TGGAATGAGT TGAGGCCACT GTGTTCTTC CGCTTGAGC 1320
205 TATTTTCCAG GCTTCTGCTG GGGCCTGGGA CAACTGCTCC CACCTCCTGA CACCCTTCTC 1380
206 CCACTCTCCT AGGCATTCTG GACCTCTGGG TTGGGATCAG GGGTAGGAAT GGAAGGATGG 1440
207 AGCATCAACA GCAGGGTGGG CTTGTGGGGC CTGGGAGGGG CAATCCTCAA ATGCGGGGTG 1500
208 GGGGCAGCAC AGGAGGGCGG CCTCCTTCTG AGCTCCTGTC CCCTGCTACA CCTATTATCC 1560
209 CAGCTGCCTA GATTCAGGGA AAGTGGGACA GCTTGTAGGG GAGGGGCTCC TTTCCATAAA 1620
210 TCCTTGATGA TTGACAACAC CCATTTTTC TTTTGCCGAC CCCAAGAGTT TTGGGAGTTG 1680
211 TAGTTAATCA TCAAGAGAAT TTGGGGCTTC CAAGTTGTTT GGGCCAAGGA CCTGAGACCT 1740
212 GAAGGGTTGA CTTTACCCAT TTGGGTGGGA GTGTTGAGCA TCTGTCCCC TTTAGATCTC 1800
213 TGAAGCCACA AATAGGATGC TTGGGAAGAC TCCTAGCTGT CTTTTTCCT CTCCACACAG 1860
214 TGCTCAAGGC CAGCTTATAG TCATATATAT CACCCAGACA TAAAGGAAAA GACACATTTT 1920
215 TTAGGAAATG TTTTAAATAA AAGAAAATTA CAAAAAATAA TTTTAAAGAC CCCTAACCT 1980
216 TTGTGTGCTC TCCATTCTGC TCCTTCCCCA TCGTTGCCCC CATTTCTGAG GTGCACTGGG 2040
217 AGGCTCCCT TCTATTTGGG GCTTGATGAC TTTCTTTTGG TAGCTGGGGC TTTGATGTTT 2100
218 CTTCCAGTGT CATTTCTCAT CCACATACCC TGACCTGGCC CCCTCAGTGT TGTACCAGA 2160
219 TCTGATTTGT AACCCACTGA GAGGACAGAG AGAAATAAGT GCCCTCTCCC ACCCTCTTCC 2220
220 TACTGGTCTC TCTATGCCTC TCTACAGTCT CGTCTCTTTT ACCCTGGCCC CTCTCCCTTG 2280
221 GGCTCTGATG AAAAATTGCT GACTGTAGCT TTGGAAGTTT AGCTCTGAGA ACCGTAGATG 2340
222 ATTTCAAGTT TAGGAAAATA AAACCCGTTG ATTACT 2376
224 (2) INFORMATION FOR SEQ ID NO: 5:
226 (i) SEQUENCE CHARACTERISTICS:
227 (A) LENGTH: 237 amino acids
228 (B) TYPE: amino acid
229 (C) STRANDEDNESS: single

```

RAW SEQUENCE LISTING

DATE: 09/26/2001

PATENT APPLICATION: US/09/938,885

TIME: 12:07:56

Input Set : N:\Crf3\RULE60\09938885.txt

Output Set: N:\CRF3\09262001\I938885.raw

```

230      (D) TOPOLOGY: linear
232      (vii) IMMEDIATE SOURCE:
233      (A) LIBRARY: GenBank
234      (B) CLONE: 945419
236      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
238 Met Ser Arg Ser Asn Arg Gln Lys Glu Tyr Lys Cys Gly Asp Leu Val
239 1      5      10      15
240 Phe Ala Lys Met Lys Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Glu
241      20      25      30
242 Met Pro Glu Ala Ala Val Lys Ser Thr Ala Asn Lys Tyr Gln Val Phe
243      35      40      45
244 Phe Phe Gly Thr His Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe
245      50      55      60
246 Pro Tyr Glu Glu Ser Lys Glu Lys Phe Gly Lys Pro Asn Lys Arg Lys
247 65      70      75      80
248 Gly Phe Ser Glu Gly Leu Trp Glu Ile Glu Asn Asn Pro Thr Val Lys
249      85      90      95
250 Ala Ser Gly Tyr Gln Ser Ser Gln Lys Lys Ser Cys Ala Ala Glu Pro
251      100     105     110
252 Glu Val Glu Pro Glu Ala His Glu Gly Asp Gly Asp Lys Lys Gly Ser
253      115     120     125
254 Ala Glu Gly Ser Ser Asp Glu Glu Gly Lys Leu Val Ile Asp Glu Pro
255      130     135     140
256 Ala Lys Glu Lys Asn Glu Lys Gly Thr Leu Lys Arg Arg Ala Gly Asp
257 145     150     155     160
258 Val Leu Glu Asp Ser Pro Lys Arg Pro Lys Glu Ser Gly Asp His Glu
259      165     170     175
260 Glu Glu Asp Lys Glu Ile Ala Ala Leu Glu Gly Glu Arg His Leu Pro
261      180     185     190
262 Val Glu Val Glu Lys Asn Ser Thr Pro Ser Glu Pro Asp Ser Gly Gln
263      195     200     205
264 Gly Pro Pro Ala Glu Glu Glu Gly Glu Glu Glu Ala Ala Lys Glu
265      210     215     220
266 Glu Ala Glu Ala Pro Gly Val Arg Asp His Glu Ser Leu
267 225     230     235

```

VERIFICATION SUMMARY

DATE: 09/26/2001

PATENT APPLICATION: US/09/938,885

TIME: 12:07:57

Input Set : N:\Crf3\RULE60\09938885.txt

Output Set: N:\CRF3\09262001\I938885.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:11 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:87 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:89 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1